

Measuring Accountability in the Summer Youth Program: A Pilot

Final Report
December 1998

Prepared by:

Andrew Wiegand
Ronald J. D'Amico
Heather Lewis-Charp

Prepared for:

Office of Policy and Research
Employment and Training Administration
U.S. Department of Labor
200 Constitution Avenue NW
Washington, D.C. 20210

CONTENTS

I. DESIGN OF THE SAMPLE	1
SELECTING SDAS TO PARTICIPATE IN THE PILOT	1
WHICH AND HOW MANY YOUTH WITHIN EACH SDA SHOULD BE SAMPLED	2
II. ASSESSMENT MEASURES	3
MEASURES OF ACCOUNTABILITY	3
AN OVERVIEW OF THE YEC SYSTEM	5
ADAPTATIONS OF YECs FOR THE SUMMER YOUTH PROGRAM	6
III. CONDUCTING THE PILOT	7
PROVIDING GUIDANCE ON ASSESSMENT PROCEDURES	7
GATHERING FEEDBACK ON THE RESULTS	7
TELEPHONE RECONNAISSANCE	7
DATA CHECKLISTS	9
CLIENT-LEVEL DATA	9
FOCUS GROUP	9
IV. FINDINGS FROM THE PILOT TEST	10
FINDINGS CONCERNING INDIVIDUAL SDAS	10
New Hampshire (Region I)10	
Middlesex County Employment and Training Department (Region II)	
13	
Philadelphia PIC (Region III)16	
Palm Beach County Workforce Board (Region IV)18	
Southeastern Minnesota (Region V)20	
Houston Job Training Partnership Council (Region VI)22	
Montana CEP (Region VIII)24	
Seattle-King County PIC (Region X)25	
ANALYSIS OF CLIENT-LEVEL DATA	27
DIFFICULTIES IN ASSESSING ACADEMIC ENRICHMENT	33
DIFFICULTIES WITH THE DURATION OF THE PROGRAM	36
THE FIT BETWEEN SDAS' GOALS AND ASSESSMENT MEASURES	38
WHOM TO INCLUDE IN AN ACCOUNTABILITY SYSTEM	40
LIMITATIONS OF THE YEC SYSTEM: ALTERNATIVES	42
CONCLUSION	46

MEASURING ACCOUNTABILITY IN THE SUMMER YOUTH PROGRAM: A PILOT

Consistent with GPRA, the Department of Labor contracted Social Policy Research Associates to conduct a pilot-test of a procedure for measuring accountability for academic enrichment in the JTPA Title II-B program. This pilot was conducted for participants in the summer, 1998 program. This final report describes the purposes for which the pilot test was conducted, the procedures and methods by which it was done, the data that were collected, and findings from our efforts.

There were several objectives of the pilot test. First, DOL was interested in determining how much effort it is for the SDA and/or its service providers to collect the additional information necessary to support the measure of accountability. Second, the pilot test sought to determine whether the accountability measures already being used capture key elements of what the SDAs are trying to accomplish and whether they serve as a meaningful yardstick of the SDAs' success. Further, the pilot test examined whether a uniform, or standardized, measure is sensible for programs with diverse programmatic objectives for academic enrichment. Finally, the pilot test investigated whether it is feasible to implement a uniform system of accountability on a broad scale. Each of these purposes was incorporated into the design of the pilot test in an effort to conduct a comprehensive investigation into the feasibility and utility of an accountability system for the Title II-B program.

DESIGN OF THE SAMPLE

Selecting SDAs to Participate in the Pilot. Nine SDAs were asked to participate in a pilot for the summer of 1998. Three alternative strategies were considered for selecting these SDAs:

Selecting purposively from SDAs that participated in a recently completed Summer Youth evaluation.

Selecting randomly within region, proportionate to each SDA's size.

Eliciting nominations from the DOL Regional Offices.

Because SDAs that participated in the Summer Youth evaluation had already been asked to bear a considerable burden, and because the information for this evaluation was collected in the summer of 1994 (and, thus, may not still apply), we decided not to select from these SDAs. Additionally, because so few SDAs are being chosen (only 9 in total), random selection had no real advantages from the standpoint of ensuring that those selected were in some sense representative of SDAs across the nation.

Accordingly, then, we asked the DOL Regional Offices to suggest nominees. These were SDAs that DOL regional staff had reason to believe would cooperate and that were known to have assessment systems for measuring performance. Given the very tight timeline within which we were operating, eliciting nominees in this fashion was thought to be an important timesaver. In light of these considerations, we suggested that SDAs be

selected by eliciting nominees from the DOL Regional Offices, following these procedures:

Nominees for at least one SDA were sought within each region. Specifically, those making the nominations were asked to choose one or more SDAs from those deemed likely to cooperate and with reasonably well-developed assessment systems.

Because some regions nominated more than one SDA, the SDA from these regions selected to participate in the pilot was drawn randomly.

Based on these guidelines, the following nine SDAs were selected to participate in the pilot test: New Hampshire (Region I), Middlesex County Employment and Training Department (Region II), Philadelphia PIC (Region III), Palm Beach County Workforce Board (Region IV), Southeastern Minnesota (Region V), Houston Job Training Partnership Council (Region VI), Montana CEP (Region VIII), City of Los Angeles (Region IX), Seattle-King County PIC (Region X). All were contacted and agreed to participate.

Unfortunately, we were unable to effectively coordinate data collection with one SDA. The City of Los Angeles agreed to participate by describing the lessons they had learned from past experience with accountability efforts. But we were unable to schedule adequate time with key contacts at this SDA and therefore could not include them in our pilot study. We wish to stress that in no way was this SDA uncooperative; instead, we had great difficulty identifying the key contact people and, once we had done this, we had further difficulty scheduling adequate time to interview them. Because we did not want to include them without having suitable time to learn important information concerning their program, we decided to exclude them from this report. As a result, the findings discussed in this report are based primarily on the remaining eight SDAs in our sample.

Which and How Many Youth Within Each SDA Should be Sampled. For purposes of developing estimates of national accountability for the pilot, not all youth within the sampled SDAs need be assessed. To minimize the burden on each SDA, youth were sampled. Logistically, a major difficulty is that Summer Youth participants within each SDA are spread out across potentially many different training sites.

¹ Thus, simple random selection would have meant that perhaps just a few participants would have been selected from each of many service providers, giving rise to the need to assess youth in many diverse locations with many different partners.

As an alternative, we used cluster sampling, whereby providers of academic enrichment were randomly selected from among all those used by each SDA, and, within the selected providers to the extent possible, at least 25 youth from each provider were assessed. Additionally, for the purposes of the pilot, we used the following criteria for selecting our sample:

¹ For example, one moderately large SDA that was included in this pilot test uses 55 separate service providers for academic enrichment, spread out over a large geographic area.

Up to 5 providers of enrichment services were selected from within each SDA; for SDAs using fewer than 5 providers, all providers were selected. The objective was to ensure that the assessment systems are evaluated against a broad selection of program activities, with potentially diverse sets of goals and objectives. Operationally, we asked the SDAs to forward lists of their providers to us, along with the approximate number of young people each served. In concert with the SDA, we made the selection of providers to participate in the pilot by including providers with as broad a range of program activities as possible.

Within each provider, we requested that up to 25 youth be tested. To avoid purposive selection on the part of the providers, we asked providers to select whole classes for assessment. For example, if a provider served several hundred young people in several different classes, youth in just one of the classes were selected for assessment.

Following DOL's guidance, only those youth receiving academic enrichment were included in this pilot assessment. Despite this, as will be noted below, incorporating participants who receive work experience into an accountability system may capture better the overall goals and achievements of the Summer Youth program

ASSESSMENT MEASURES

Measures of Accountability. Measures of accountability for the Summer Youth program must have several features in order to be consistent with the purposes of the program itself. For example, these measures should provide indicators of accountability that reflect goals and objectives that SDAs are already pursuing in the context of academic enrichment in the Summer Youth program. Similarly, they should be wary of pushing SDAs towards any particular service design or set of services. The Department of Labor has deliberately allowed SDAs great flexibility in pursuing academic enrichment activities that best meet the needs of their communities. This flexibility is an important component of the Summer Youth program, not an aspect which should be lost in an effort to steer all participants toward a uniform service design. Consistent with this, the measures must not give rise to any perverse incentives. Past experience has shown that measures of accountability can cause programs to change their behaviors in ways that are unintended and even undesired. In so far as possible, such perverse incentives need to be anticipated and avoided.

Further, the measures should have reasonable properties with respect to validity and reliability. In other words, they should reflect true measures of accomplishment and be measured reasonably consistently from grantee to grantee. Without reliability in measurement characteristics, data collected from grantees cannot be taken to indicate the same information. Thus, what is defined as a "success" by one grantee may be entirely dissimilar to what is defined as a "success" by another.

Additionally, accountability measures should be relatively easy to implement. Burdensome procedures put in place in an effort to provide accountability may serve only to detract from the true purposes of the program. As a corollary, these measures should impose only modest additional burden on SDAs and their providers. Summer Youth

programs are already hard pressed to enroll, assess, and serve thousands of young people within a very compressed timeframe. The accountability system should not add unduly to this burden.

Based on these considerations, DOL suggested adopting Title II-C Youth Employment Competencies (YECs) as the accountability system to be piloted this summer. There were several reasons for this choice. First, the YEC system was already very familiar to SDAs. Additionally, YECs allow SDAs substantial flexibility in carrying out their summer programs; as such, they can accommodate diverse goals and objectives, without forcing SDAs explicitly or unintentionally to revamp their service designs. These advantages suggested that YECs best met the criteria outlined above, at least given the brief timeframe within which the pilot test was conducted. Meanwhile, we simultaneously explored ways in which the YEC system can be embellished or modified, to enhance its rigor in reflecting meaningful program accomplishments. The outline of this plan is described below.

An Overview of the YEC System. As mentioned, using YECs as the foundation for the pilot test offered several advantages. A brief description of YECs, and how they are currently used for assessment in the Title II-C program, is presented below.

The purpose of YECs is to measure and assess youth competencies in the following skill areas:

Pre-employment skills and work maturity (PE/WM) skills. Eleven core PE/WM competencies have been identified, including:

- Making career decisions
- Using labor market information
- Preparing resumes
- Filling out applications
- Interviewing
- Being consistently punctual
- Maintaining regular attendance
- Demonstrating positive attitudes/behaviors
- Presenting appropriate appearance
- Exhibiting good interpersonal relations
- Completing tasks effectively

Basic education skills. These include primarily reading comprehension and math computation, but also writing, speaking, listening, problem solving, reasoning, and the capacity to use these skills in the workplace.

Job-specific skills. These encompass the proficiency to perform actual tasks and technical functions required by certain occupations at entry.

In support of performance standards in the Title II-C program, the Department has issued minimal structural and procedural elements of a “sufficiently developed youth employment competency system.” These include the following key components.²

Quantifiable learning objectives. The SDA should establish quantifiable learning objectives related to the YECs that include a description of the skills, knowledge, attitudes, or behaviors to be taught, the levels of achievement to be attained, and the means of measurement to be used to demonstrate competency accomplishment.

Training modules and approaches. The SDA should identify the curricula, components, or courses that will teach the competencies and should include, as appropriate, relevant manuals, implementation packages, instructions, and guidelines.

Pre-assessment. The system should include the assessment of participants’ employment competency needs at the start of the program. A minimum level of need must be established before a participant is eligible to be tracked as having attained a competency in one of the above areas. All assessment techniques must be objective, unbiased, and conform to widely accepted measurement criteria.

Employability development planning. The pre-assessment results should be used in assigning a youth to appropriate learning activities.

Post-assessment (evaluation). The SDA must evaluate the participants’ achievements by the end of the program to determine if competency-based learning gains took place during project enrollment. As with the pre-assessment, all evaluation techniques must be objective, unbiased, and conform to widely accepted measurement criteria.

Documentation. The SDA is expected to maintain participant records to document any competency-based learning gains that are claimed.

In the Title II-C program, a participant can be claimed as having attained an employability enhancement, if he/she attains competencies in at least 2 of the 3 areas, defined as PE/WM skills, basic skills, or job-specific skills. To claim an attainment of a PE/WM competency, at least 5 of the 11 skill areas identified above must have been achieved; typically, attainment of basic skills competencies is claimed for participants who increase their reading or math skills on a standardized test by a specified amount or who achieve a certain level of proficiency.

Adaptations of YECs for the Summer Youth Program. Rather than impose a uniform assessment system on each of the SDAs participating in the pilot program, we wanted to take full advantage of the opportunity to learn about diverse assessment systems to better gauge what measures work best to document the achievements of Summer Youth programs. As such, we agreed to use in the pilot test those measures already in place for each of the SDAs. These measures are primarily adaptations of the YEC assessment system, modified to fit better the goals of each program. By using those measures already in place, we are able to assess the utility of a wide range of measures, rather than constrain each SDA into a single mold. Although this can create other problems, such as

² These guidelines are lifted almost verbatim from the Department’s issuances.

a lack of standardization across all SDAs, because each of the participating SDAs uses modified YECs there are similarities in assessments across programs which allows us to compare the relative utility of the various measures.

CONDUCTING THE PILOT

The several stages of the pilot test included: selecting the SDAs and participants to be assessed, providing guidance to SDAs on the assessment procedures, collecting a variety of data from SDAs and their providers, and gathering feedback on the results of the pilot. The procedures for selecting the SDAs and participants have been discussed above. The remaining steps associated with the pilot are sketched briefly below.

Providing Guidance on Assessment Procedures. Once SDAs and some of their providers were selected to participate in the pilot, we developed protocols for telephone interviews which were designed to determine what measures of accountability were already in place. We took it as a point of departure that each SDA was familiar with YEC concepts and had “sufficiently developed” assessment systems in place. Thus, the protocols were structured to determine how each SDA had adapted YEC concepts to better fit the goals of their program. By determining what assessment instruments were being used, we could then provide assistance to the SDA and its providers in how best to sample participants, and the types of data needed for the pilot.

These conversations also allowed us to gain important background information on what were the SDAs’ goals and objectives for academic enrichment, what methods they used to achieve them, and how they communicated these goals to their providers. Such information allows us to provide a more detailed analysis of what accountability measures are best suited for particular goals.

Gathering Feedback on the Results. In order to meet the objectives of the pilot, we needed to know how well the accountability measures captured the accomplishments of the programs being assessed and how easy they were to implement. For this purpose, we conducted four data collection activities: telephone reconnaissance with the SDA’s Summer Youth Coordinators and the service providers participating in the pilot, having providers complete checklists detailing how many youth were assessed and how many achieved their competencies, collecting more detailed, client-level data with pre- and post-test results for many participants, and conducting a focus group with Summer Youth Coordinators from non-pilot sites.

Telephone Reconnaissance. At the conclusion of the summer, we conducted initial telephone conversations with Summer Youth Coordinators and the service providers who participated in the pilot. These conversations followed loose scripts that ensured that we systematically covered the following key points, among others:

- Did their accountability system capture what respondents’ view as their programs’ chief accomplishments? If not, what types of learning gains were they hoping to impart that were not captured by the assessment?
- More generally, what were their programs’ chief objectives?

Conversely, were elements of the accountability system inappropriate for their programs' goals and objectives?

What competencies were established for youth in the programs and how were attainments assessed and documented?

How much effort was it to conduct the necessary assessment, given the practical constraints associated with the Summer Youth program (e.g., short timeframes, youth who may be shuttling between worksites and classroom activities).

What types of instruction are offered?

Did the SDA choose the assessment measures? Does the provider have flexibility to modify the measures to better fit their program?

What changes would be made to capture better the achievements of youth in the program?

Additional questions were directed only at the Summer Youth Coordinators, including:

How does the SDA interpret the definition of academic enrichment? What range of activities does the SDA and its providers conduct under this rubric?

What are the elements of the SDAs accountability system?

How feasible would it be to assess employment competencies for all youth receiving academic enrichment?

How feasible would it be to implement a similar assessment system for youth receiving only work experience? What would be the challenges in doing so? Could the learning objectives associated with work experience positions be assessed in this way?

As mentioned above, these phone calls were completed for eight of the nine SDAs in our sample, and for many of their providers.

Data Checklists. As additional output for evaluating the success of the pilot, we prepared checklists to be completed by those conducting the pilot assessment, enabling them to record: the number of youth who were served, the number who were assessed, the number who were determined to have achieved their competencies, and (when available) the number who achieved in each of the competency areas (e.g., each of the PE/WM skill areas, basic education, and job-specific skills). Nearly 70% of pilot SDAs and their providers returned these checklists to us, giving us a context within which to frame the results of this study.

Client-Level Data. We also asked SDAs and their providers to send client-level data for their participants. Such data allow us to analyze the gains made by participants on an individual basis, offering a much more powerful tool for understanding the success of the Summer Youth programs, and also how well a particular measure documents

performance. Further, these data provide a wealth of information about the participants, allowing us to examine in greater detail the performance of SDAs and their providers. Analysis of these data enables us to examine what types of participants achieve particular types of gains in the Summer Youth program, as well as what types of program instruction or emphases lead to such gains. The results for this analysis, presented below, are described across all sites reporting client-level data, rather than for any single SDA or provider.

Focus Group. In conjunction with the Summer Youth training conference conducted in Chicago on October 7-8, we conducted a focus group with some of the attendees. This focus group offered the chance for a wider range of SDAs (specifically, those not participating in the pilot) to express their thoughts on the utility of various assessment systems for the Summer Youth program. Issues discussed included some of those already mentioned above, such as:

How do SDAs interpret the meaning of academic enrichment?

What are the goals and objectives of their academic enrichment programs?

What specific skills are being imparted?

Does their assessment system capture the programs' chief achievements? If not, why not? What is being missed?

How practical would it be to implement assessment on a broad scale for Title II-B? How practical would it be for those in work experience?

FINDINGS FROM THE PILOT TEST

We begin by summarizing briefly some of the findings from our initial interviews with SDAs and service providers. As mentioned, several rounds of telephone interviews were conducted, including many of those questioning providers as to how easy it was to collect the information necessary for our pilot test, and whether they believe the results adequately capture the performance achievements of their program. Responses to these and other questions are incorporated into the following descriptions of each of the eight SDAs whom we were able to interview. These descriptions summarize what assessments instruments each SDA already had in place, which were used for this pilot test, and provide an indication of whether or not these measures adequately reflect the overall programmatic objectives of each SDA.

Findings Concerning Individual SDAs.

As mentioned above, originally there were nine SDAs in our sample. Unfortunately, we could not effectively coordinate data collection with one of these. Below, therefore, we present findings from the remaining eight SDAs.

New Hampshire (Region I). This SDA serves about 800 youth through the Summer Youth program, and these youth are served by approximately 40 providers. As can be seen in these numbers, many of the programs are quite small, serving 10 to 15 youth. This allows the programs to be quite intensive, and have a large amount of individual contact with the youth. Indeed, one extremely positive feature of this design, according to

the SDA, is that the ratio of instructors to participants is often 1:10 and, in some cases, 1:5. The SDA believes keeping the ratio low offers the best opportunity to reach, and teach, at-risk youth. The youth served by this SDA generally fall into the category of at-risk, as they are predominantly out-of-school. Many participants in the summer program recruited through their involvement with the year-round JTPA Title II-C programs.

About 75% of the youth served by this SDA during the pilot test were 14-15 years old, and through the Summer Youth program were receiving their first job. The primary reason for such a high percentage of participants in this program being so young is that New Hampshire's unemployment rate was very low. As a result, older youth were able to find a job on their own, and often these jobs pay more than the minimum wage paid in jobs obtained through the New Hampshire Summer Youth program. As this explanation would imply, this proportion of younger participants has fluctuated dramatically in New Hampshire as the economy has flourished or declined.

These changes, obviously, make it difficult to offer a staple set of activities or options, because as the population served changes, so too does the need for various services. And as these services change, we would expect that the areas in which participants demonstrate the greatest improvement or gains also would change. Thus, we should not interpret the results from our pilot test to be easily generalizable to all participants in the Summer Youth program in New Hampshire, regardless of the economic circumstances in the state.

Providers are responsible for recruiting at least 75% of their participants. In an effort to ensure that providers do not inflate their success by choosing only the most promising candidates, the SDA requires each provider to rank their applicants based on a number of potential barriers to success. Those who score the highest, indicating they have the greatest number of barriers to success, must be accepted into the program. As a result, simply measuring the number or percentage of participants who demonstrate gains in competencies over the course of the summer program may not fully capture the success of each program or provider, because some participants face greater obstacles to success than do others.

Each participant in the New Hampshire SDA's Summer Youth program receives a combination of academic enrichment and work experience. Participants receive a minimum of 90 hours of academic enrichment, and a maximum of 120 hours of work experience. The program ranges from 6 to 9 weeks, depending on the provider.

This SDA defines academic enrichment quite broadly, as they feel DOL intends it to be interpreted. A primary goal of this SDA for academic enrichment is that it should work toward an integration of work activity and project-based learning, or an integration of work and school environments. As an example, many of this SDAs' providers ask students to build a portfolio throughout the program. Where this is done, the portfolio generally consists of a completed employment application, resume, cover letter, references, and samples of their work. This demonstrates that the items on which the SDA evaluates are actively attempting to integrate academic skills (i.e. reading, writing) with the world of work (i.e. resumes, etc.).

The overall goal of the New Hampshire SDA is twofold. First, they hope to provide youth with a positive work experience. In this way, the youth will learn to associate work with a pleasant experience, rather than a meaningless chore. Second, the SDA hopes that participants leave the program at the end of the summer with the same or better academic skills than they had when they entered the program.

This second goal echoes the theme mentioned above: the SDA does not set specific benchmarks for *gains* in these skills. This is intentional for at least two reasons. First, the Summer youth program is short in duration, generally lasting only 8 weeks. It is somewhat unrealistic to expect to see meaningful increases in academic skills in such a short program, only part of which involves training in these basic skills. Second, however, is that the participants in these programs generally have significant obstacles to academic success. Because of this fact, it is even more unreasonable to expect that gains in academic skills would be observed as a result of this program. Some participants do, of course, demonstrate such gains. But the SDA clearly believes that helping the participants to maintain their academic skills, instead of diminishing over the course of the summer, is a legitimate goal of their program.

To evaluate their success, this SDA and its providers use the Comprehensive Adult Student Assessment System (CASAS) for a pre- and post-test assessment of both academic/basic skills and pre-employment/work maturity skills. They have been using the CASAS for over 10 years. Additionally, some providers for this SDA use other, supplementary assessments, such as the ACT for college-preparatory providers. The results from these tests are compiled for each student into a report card for the summer. The student receives one copy of this report card, one copy remains with the provider in whose program the student participated, and one copy is given to the SDA and remains in their client files.

The SDA allows providers to select certain portions of the CASAS most relevant for their programs. For example, those providers emphasizing basic skills and GED attainment can select those sections of the CASAS most relevant to this emphasis. Similarly, those providers emphasizing work readiness can select portions of the CASAS designed to assess these variables. To the extent that providers do tailor the use of the CASAS to their needs, they report it being a useful measure for evaluating participants. In fact, several providers indicated that they use results from the CASAS pre-test to help to design an individual service plan for their participants. Others report that they structure their classroom instruction around issues which will be assessed using the CASAS. In this way, they can be certain they are providing instruction in the areas emphasized by the SDA.

The SDA also produces a report card for each provider, documenting their performance in attendance, the percentage of students who reach competencies, and the percentage of students who maintained the same level (or improved) in their basic skills. This allows the SDA to examine areas of strength and weakness for each provider. Additionally, this report card allows the provider to have a written documentation of their performance.

Finally, the SDA also conducts a customer satisfaction survey each year in an effort to determine how the participants themselves rate the program.

The assessment instruments of this SDA, then, match quite closely their goals for the program. The CASAS provides an assessment of the participant's skills, which allows the SDA to determine whether the participant maintained, or even improved their skills during the course of the summer. The thrust of the CASAS evaluation for this SDA is its focus on the world of work and work themes. The SDA reports that the CASAS provides a generally reliable and thorough assessment of its performance, and that the providers generally favor it, as well. By focusing on the world of work, and integrating basic educational skills' assessment into work themes, the CASAS allows the SDA to meet its goal of integrating school and work contexts, while evaluating the participants' performance in the Summer Youth program.

Further, results from the customer satisfaction survey allow the SDA to determine whether it is achieving its first goal: providing the participants with a positive work experience. Thus, the instruments used by this SDA closely parallel their two primary programmatic objectives, allowing them to determine whether their efforts were successful.

Additionally, many providers use their own assessment instruments, including portfolios, in an effort to better evaluate achievement toward their own goals. While these may be less essential to the SDA, they do provide a broader picture of success in the varieties of training and experience offered through the New Hampshire SDA.

Middlesex County Employment and Training Department, New Jersey (Region II). This SDA served slightly more than 900 participants through 16 providers in three counties in New Jersey. Two of the providers were community-based organizations, 13 were secondary or post-secondary schools, and 1 was the Central Labor Council. The SDA chooses providers after examining the curriculum offered. The SDA attempts to choose those curricula which emphasize pre-employment/work maturity skills.

Recruiting is done by both the SDA and the schools. While no specific age group is targeted, the majority of students in this SDA's program are between 14 and 16 years old, because these are the individuals who have the greatest trouble getting jobs on their own. A strong majority of the participants in this program are in-school youth. This fact makes it easier for the SDA to focus on pre-employment/work maturity skills, because the participants are receiving training in basic skills during the traditional academic year.

36 of the 900 participants received work experience only. These were primarily older participants. The remainder of the program's participants were involved in both academic enrichment and work experience. The program focuses on integrating basic skills into "real life" or work-based situations. These situations include projects, work maturity skills (e.g., interviewing, resume writing) and life skill development (e.g., opening a bank account, comparison shopping for the best prices). The programs range in duration, depending on the particular provider, but are generally between 6 and 8 weeks long.

Academic enrichment is defined by this SDA as any “special event” which gives youth a first-hand experience of the work world. Some events which qualify under this label are speakers who talk to youth about the world of work, and field trips which take youth to visit work sites. Thus, traditional classroom instruction does not fit within the definition of academic enrichment used by this SDA, although a significant percentage of their time is devoted to classroom-based instruction. Much of this classroom instruction, however, is not geared toward improving basic skills, as noted above. As a result, no specific benchmarks are set for educational attainment during the summer. According to the SDA, if youth remain at the same level, or show minimal improvements, in their academic skills by the end of the summer, they are considered a success for the program. This relatively modest expectation is a result of the short duration of the summer program.

Rather than increases in academic skills, then, the overall goal of this SDA’s Summer Youth program is to help youth develop a strong work ethic. They would like to help participants understand what it means to show up for work on time, dress appropriately, earn their own income, and interview for a job, among other skills. An additional goal is to introduce youth to a variety of people and possible work situations, or, in other words, to increase the social resources of these participants who generally do not have great access to employers and employment situations.

This SDA has been using the YEC assessment system, both for its year-round (Title II-C), and summer (Title II-B) participants. For the summer program, the YECs were designed to assess the impact on participants’ basic skills and pre-employment/work maturity. The New Jersey SDA allows each provider to choose a suitable basic skills measure from a list of acceptable tests (including many of those used by other SDAs in our sample, such as the TABE, and ABLE). Additionally, a pre-employment/work maturity competency instrument was given to all participants. Each youth received a pre- and a post-test for both these instruments in an attempt to document any gains made during the Summer Youth program.

The pre-employment/work maturity measures allow the SDA to document performance in terms of their first goal for the program. These measures are employer ratings, allowing each employer to rate their youth's work maturity skills both at the beginning of the work experience, and toward the end of the program. By requiring employers to evaluate participants on these skills both before and after their participation, the SDA can assess the extent to which participants’ work maturity skills increased as a result of their participation in the Summer Youth program. Again, however, the SDA does not set specific benchmarks for these measures, believing that it is unrealistic to set such standards for a brief summer program. Indeed, one representative from this SDA asserted it was too much to expect that this summer experience would play a valuable role in helping youth to choose a career path because many participants are too young, and the program is simply too short. However, with the assessment instruments in place, the SDA can determine whether their program is increasing the work maturity skills of its youth.

In contrast, the basic skills measures in place do little to document performance in a

meaningful way for this SDA. This is because the SDA believes that it is unlikely to achieve gains in basic skills in such a short period of time, especially when basic skills are only one component of the program. Thus, these measures are in place, but little weight is placed on observing gains in youths' skills. Instead, to the extent that youth maintain the level of skills they brought into the summer program, the program is considered a success.

Because the SDA uses a basic skills measure, but does not expect to see gains among their youth due to the short duration of the program, there is a feeling among some at the SDA, and some providers, that the assessment measures are too broad for the Summer Youth program. This feeling is compounded by the belief that assessment guidelines are not well-specified and, as a result, the SDA is unsure how best to capture the gains their participants have made during the program. To the extent that a national accountability system is well defined, and put in place well in advance of the beginning of the summer program, then, this SDA believes their participants' progress could be better captured.

Philadelphia PIC (Region III). This SDA served over 6,200 youth (out of nearly 25,000 applications) through 55 separate providers. Of these, 1,600 received classroom-only instruction and the remainder were either primarily in work experience or in the service corps, both sets of which received some combination of work experience and academic enrichment.

Recruitment for the program primarily is conducted by the SDA, although 15 of the providers do their own recruiting because they target distinct populations of youth (i.e., disabled youth). Although there is an initial test of basic skills, the results of this assessment are not used in determining eligibility or choosing those youth who will participate in the program. The recruitment takes place generally through the school districts, indicating that the majority of youth in this program are in-school youth.

Providers are chosen based on their curricula. The SDA identifies those curricula which attempt to use work-based learning and incorporate school-to-work principals. Specifically, the SDA chooses curricula which form linkages between classrooms and the workplace.

Academic enrichment is defined by this SDA to be anything which focuses on the “next step.” In other words, anything which helps participants to understand the workplace, their place in it, and the importance of education in reaching their goals regarding the world of work is considered by this SDA to be enrichment. Clearly, this definition is broad, and can incorporate a wide variety of concepts and training. As such, it can be difficult to assess the broad range of training offered by the various providers for this SDA.

The overall goal of the Philadelphia SDA is to teach participants excellent work skills, and the importance of education in reaching their work goals. To this end, the SDA attempts to monitor the curriculum of each provider, as well as monitor the grades and attendance of each participant. Additionally, the SDA provides a wealth of technical assistance to its vendors, including three case managers who make regular site visits to

the provider, or to employers, in an effort to monitor the quality of the program.

For formal assessment, this SDA has been using a variant of the YEC assessment system, with a pre- and post-test using the MAST measure to assess basic skills. In particular, each participant is given assessments in basic math and reading skills to determine their grade level in each at the beginning and end of the program. Additionally, they use a work maturity measure completed by the participants' worksite supervisor. Each provider forwards the results from these measures to the SDA.

The SDA also monitors their participants' performance in school throughout the following academic year in an effort to determine whether participants have better attendance, or receive higher grades, than before their Summer Youth experience. This represents a significant amount of effort, but because the overall goal of the program is to help youth see a connection between school and their future careers, such long-term monitoring is important. The reason for this seems clear: to the extent that youth understand that education is important to secure desirable employment in the future, they should remain in school in greater rates than do youth who do not understand this connection.

The SDA has established fairly strict benchmarks for their providers and participants. For example, the SDA expects that participants will receive at least satisfactory ratings in 8 out of the 11 work maturity categories. Also, participants are expected to gain at least 1 grade level in their basic skills, as determined by their post-program MAST score. Further, the SDA expects youth to receive at least 70%, or a grade of "C," for their academic enrichment coursework throughout the summer in order for the youth to be considered a successful termination.

From these benchmarks, it is clear that the SDA emphasizes academic skills, and attempts to increase these skills in its participants. Indeed, many participants are able to earn an academic credit toward their graduation through their participation in the program. Compared to other SDAs, then, the Philadelphia PIC clearly stresses the increase of basic skills far more heavily. This may be due to the population served by the program. Specifically, most participants are in-school youth, but many face obstacles to high achievement in their schools. The Summer Youth program attempts to overcome some of these barriers, and allow the participants to increase their academic skills.

Each of these benchmarks used for the providers, then, stress the importance of academics in the summer program. But the benchmarks, and assessments, do not seem entirely related to the overall goals of the program. For example, the MAST scores do not indicate whether the participant has developed excellent work skills, nor whether the participant understands the importance of education in achieving their work goals. Instead, these scores simply indicate the education level of a particular participant. Perhaps increases in academic scores indicate a greater perception of the interconnectedness of school and work, but this is not certain. As a result, these benchmarks, while impressive, do not by themselves document whether the SDA is meeting its goals. In contrast, the work maturity component does offer a rating of the participants' work skills, and can help to document whether providers are helping to

improve their participants' work readiness. This measure, therefore, is well-suited to the goals of the SDA.

Palm Beach County Workforce Board (Region IV). This SDA serves slightly more than 1,500 youth. These youth are served by four separate providers; each of these providers operates their own program, and the programs vary greatly. As a result, the class size for participants varies a great deal, as well.

The participants in the program also vary in age, from 14 to about 21. Most of the younger youth (ages 14-17) receive primarily academic enrichment. These youth are in-school year-round, and the basic skills training they receive during the summer is a supplement to their year-round training. Graduating seniors often receive a combination of basic skills remediation training and work experience. Older youth (18-21) who are out-of-school (with or without their high school diploma or GED) receive primarily work experience. The program served about 200 18-21 year olds, and the remainder were 14-17 year old in-school youth. The program duration ranges from six to eight weeks.

Out-of-school youth are given two days of training in how to interview for a job, how to dress, appropriate workplace attitudes, and other work-related knowledge. This is considered by the SDA to be academic enrichment. After this brief session, however, the youth are given a list of potential employers, they set up interviews, and upon accepting a job, work for seven weeks. Following this experience, they receive an evaluation which can serve as a reference for future work.

In-school youth, in contrast, receive predominantly academic enrichment. The youth spend four days per week (for six weeks) in classroom-based instruction. This training focuses on developing the youths' math and reading skills in a work-based curriculum. On the fifth day of each week, youth work at a work site set up by the provider to help the youth integrate their classroom training with the world of work. Class size for these youth is kept to a minimum, about 12 students per class. Upon successfully completing the program, the youth receive an academic credit toward their high school graduation.

A third provider offers what is described as the "Next Step" program. In this program, a community college provides both academic enrichment and remedial training for graduating seniors. Specifically, the college focuses on developing professional skills, offers a class on college preparation, and places youth in worksites. Academic training and college preparation is done in the morning of each day during the program, and work experience is gained during the afternoon. Youth in this program receive \$5.50 per hour for their work, and receive college credit for the courses they take.

A fourth provider serves a much smaller population of African-American youth. This provider, too, offers youth classroom instruction in the morning, and work experience during the afternoon, but it focuses specifically on African-American communities, literature, and culture.

According to this SDA, academic enrichment is a very broad concept. Included in this concept are remedial math, reading, and communication skill development, as well as

work place skills (e.g., dressing for an interview), and entrepreneurial skills (e.g., instruction in basic economics).

The overall goal of this program is to help participants develop skills which will help them function in life. The SDA believes that although the participants generally have not thrived in academic settings, they will appreciate the career-based instruction offered during the summer and, therefore, perform better in this academic setting. Also, the SDA hopes this renewed interest in academics will carry over into their traditional schooling. As mentioned, the vast majority of the youth served by this SDA are in-school, and several are contemplating college. Thus, the SDA hopes to supplement the skills these youth already have while making them more interested in, and excited about, their education.

This SDA has been using the Test of Adult Basic Education (TABE) as an assessment measure for each provider, although as described it has a very diverse set of providers. The TABE has been given until recently only as a pre-test to identify deficiencies. For the pilot study, however, the TABE was given to participants as both a pre- and a post-test in an effort to assess learning gains among their participants.

To the extent that the goals of the Palm Beach SDA include academic training (clearly these are skills which help the participant function in life), then the assessment instrument in place is very well suited for documenting performance. Because the vast majority of students primarily receive academic training, albeit in a career-based curriculum, the skills most likely affected by their summer participation are their academic skills. As such, the post-test TABE results should show gains in, or at least the maintenance of the participants' original level of, academic skills.

There is, however, little documentation of participants' work-readiness skills. Because work preparedness is certainly a skill which can help the participants function in life, this lack of a work-readiness skills measure therefore does not allow this SDA to assess their participants' performance on one of their important programmatic objectives.

In sum, then, the SDA has measures in place which can help them to assess several of their overall objectives. The use of basic skills measures certainly provides an evaluation of the effectiveness of the providers' academic training. But there is little in place to assess the value or benefit of the work experience offered to participants. Some at the SDA acknowledge this when they assert that the use of the TABE helps to assess the extent of summer learning loss, but does little more for their program. Additionally, the TABE is problematic in that it can only be administered to a relatively small number of youth at a time. Thus, the test administration, while only assessing one area of the program's emphasis, also requires a great deal of time.

Southeastern Minnesota SDA (Region V). This SDA serves 250 participants in its summer program, covering a ten county area. The SDA serves as its own provider. According to the SDA, this simplifies their program a great deal, in part because they are in charge of all activities offered to their participants. Additionally, the SDA operates a year-round alternative high school, from which they borrow much of their staff for the summer program. This helps to increase the continuity of the program's offerings,

especially because many of the year-round participants also enroll in the Summer Youth program. This also helps the SDA toward achieving one of their goals, which is to provide role model for the participants.

Of the 250 participants, 15% receive only work experience, 20% receive academic enrichment only, and 65% receive some combination of work experience and academic enrichment. The program targets 14 to 15 year old participants, because these individuals have the most difficulty finding jobs. Also, they target those individuals who are having trouble making the transition from junior high to high school. As mentioned, the SDA has an alternative high school as part of their Title II-C program, and many participants in the summer program are recruited from this school. Because of the nature of the population they serve, the SDA attempts to provide academic training in non-traditional ways. Because many of the participants in the summer program are drawn from the alternative school, it seems clear that they have not performed well historically in a traditional educational environment.

Academic enrichment in this program is defined as life- and work-skill development. Specifically, the SDA attempts to incorporate math and reading skills into the work place. Generally, they try to make their academic and work activities more integrated and less compartmentalized. Although some SDAs focus more on academics in their enrichment activities, this SDA believes that enrichment primarily is about team building, group activities, and understanding how academic skills fit into the world of work.

The overall goals of this SDA are to increase the participants' opportunities to be successful in life. The top priority is for participants to leave the program with a feeling of "I can." Further, the SDA hopes that participants are affected by the program, and that through the program they develop a dream and the desire to follow it. By keeping the instructor/work site supervisor to participant ratio very low (i.e., 4:1), the SDA believes it offers its participants opportunities to achieve these goals.

This SDA already has been using a variant of the YEC assessment system for the Summer Youth program. Specifically, they use the Adult Basic Learning Exam (ABLE) to assess participants' basic skills. Additionally, they also measure pre-employment/work maturity skills, as well as job-specific skills. Because many of the summer program's participants are recruited from the year-round Title II-C program (which focuses on basic skills and pre-employment/work maturity skills) the summer program focuses primarily on the measure of what they consider to be job-specific skills. This includes some rudimentary job-specific skills, such as daycare or clerical work, but also includes some skills which might best be considered work readiness. For example, showing up to work on time and the ability to interact successfully with one's supervisor are considered job-specific skills in this definition. Finally, this SDA uses an instrument intended to capture 40 developmental assets, including interpersonal skills and ratings of self-confidence; the participants are asked their own attitudes about these assets.

Many of the measures in place seem ill-suited for the overall goals of the Minnesota SDA. Specifically, the basic skills measures seem largely unrelated to the concept of

finding a dream. This belief was confirmed by the SDA's Summer Youth Director, who feels that the Developmental Assets index is the most applicable measure for their program. By identifying what assets the youth believe they have, and assessing participants' self-ratings on a variety of personal or interpersonal measures, the Summer Youth Director believes the SDA would be serving youth better than by simply trying to assess increases in their basic skills or work readiness skills.

Houston Job Training Partnership Council (Region VI). This SDA served nearly 6,000 youth this summer. Of these, nearly 3,700 received academic enrichment only, slightly less than 2,000 received work experience only, and the final 292 received a combination of work experience and academic enrichment. Older participants are generally placed into work experience, while those participants ages 14 to 16 are placed into academic enrichment (partially due to age laws for employment in Houston). The program lasts for six weeks.

The participants who received both enrichment and work experience were given one week of training concerning how to get a job, as well as how to keep a job. They then were put into a worksite for a five week job placement. Ultimately, according to some at the SDA, this is what they would like to offer all participants in their program. However, current staffing and space limitations prevent this from becoming a reality.

Providers are chosen based on their curriculum. The SDA examines each potential curricula and chooses those which "don't look like school." The reason for this is that the participants in the program have not done well in traditional school historically and, as a result, often are alienated from programs which resemble a traditional school environment. There is a great deal of continuity among vendors, though, as many have more than ten years of experience working with the SDA. Providers also are encouraged to develop similar programs and services. The SDA does this in an effort to ensure that if a participant moves, or must transfer to another provider, they can enroll with the new provider, and still feel some continuity in the program.

Prior to the current administration, academic enrichment at this SDA was primarily classroom instruction in work maturity skills, life skills, and multicultural diversity. Since the early 1990s, however, this SDA has focused on making enrichment activities more work-based. Currently, they structure work-based learning into projects for the participants, and attempt to put academic instruction into group projects, such as by having youth develop a newsletter. The evaluation of participants is based predominantly on improving the overall group project, rather than any one individual.

The overall goal of the program is to help youth secure a job, and be able to understand what is required to maintain their employment. A success for this SDA is when a participant can secure their own job the following summer with no assistance from the Summer Youth program or this SDA.

As part of the pilot study, Houston added additional assessments to those already in place. Until recently, participants have been given the TABE and the PDI as pre-tests only. Until this year, they had found that post-testing was not a financially viable activity, because the goal of the SDA was to serve as many youth as possible. This year, however,

as part of the pilot test, participants also were given these two assessment instruments as a post-test to assess learning gains among their participants.

Because the goals of this SDA are to help youth secure a job, and to be able to secure employment by themselves the following summer, their assessment instruments are not well suited for their program. While work readiness skills may be one component of successfully securing a job, it is not the only one. As a result, this SDA has little way of documenting success rates for their participants in terms of what they want to be providing for these participants. An assessment of the percentage of participants who successfully secure employment the following summer may be a better indication of the success of this SDA. The program does, however, have some measures of work readiness and basic skills, both of which certainly can increase youths' probability of securing a job on their own.

Additionally, the SDA greatly values portfolios as an assessment instrument. Using these portfolios, the SDA can observe that the participants can complete an application, produce a resume, and write a cover letter, among other things. These portfolios clearly are related to the overall goals of the SDA: helping participants to secure employment the following summer. A portfolio may be the instrument most likely to assess a participant's future employability. As a result, then, the SDA does have in place a measure which meets their programmatic objectives. The utility of portfolios for a national accountability system will be discussed below.

Montana CEP (Region VIII). This SDA serves approximately 200 participants. Up to 60% of these participants also are involved in the year-round Title II-C program. Although a few of these are involved in basic skills remediation only (often based on age-grade-level appropriateness), the majority of participants receive a combination of academic enrichment and work experience. Because a majority of participants are recruited from the Title II-C program, which focuses on basic skills and work maturity skills, the summer program is often used to focus on job-specific skills. As with the Minnesota SDA, this definition incorporates some rudimentary job-specific skills (e.g., lawn maintenance, or database management), as well as some work readiness skills, such as showing up to work on time and being able to interact with one's supervisor.

This SDA is somewhat unique because of its small size, and the rural nature of many of their placements. As a result, their comparability to other SDAs in terms of their emphases and assessments may be limited.

A few participants in the program are involved in basic skills remediation. Youth are placed into such remediation based in large part on results of a pre-test. Those who score significantly below their appropriate age-level are eligible for remediation. The primary goal for these participants is to prevent, to the extent possible, summer learning loss. There is little expectation that the youth in remediation will demonstrate gains in their basic skills.

Academic enrichment is defined by this SDA very broadly. The SDA purposefully defines enrichment broadly because there is a great diversity of activities offered as part of the Summer Youth program. Thus, academic enrichment can include safety

instruction, presentations by professionals in the community, resume skill development, and many other activities, as well. The overall goal of the program is to provide the participants with work experience, and to help them understand how their basic skills can be applied in a work setting.

This SDA uses the YEC system to assess summer youth participants in much the same way as they use it for Title II-C programs. They use the TABE as both a pre- and a post-test to assess the impact of the program. Also, employers complete evaluations of participants' work site skills. This is done continually throughout the duration of the summer in an effort to constantly monitor the progress of the program's participants. The definition of when a youth has achieved a competency in a particular area, as well as the evaluation of participants' initial work skills, is largely left up to the work site supervisor.

Because this SDA's goals are to provide youth with job experience and to help them understand how their skills relate to the work world, their assessment instruments seem somewhat related to their goals. If, for example, worksite supervisors rate participants based on how well they understood that a particular need required a particular skill, then the assessments would indicate whether participants saw the connection between skills and the work world. But because work site supervisors have tremendous amounts of discretion in determining what, exactly, defines a competency, it is unclear that there can be any uniform assessment across participants. Therefore, the instruments currently in place seem ill-suited for assessing the overall performance of the SDA on its primary goals. As will be noted below, this assertion can be made regarding the use of YECs, generally.

Seattle-King County PIC (Region X). This SDA serves 1,600 participants through 5 different providers. Of the participants, 35% receive only work experience, 10% receive academic enrichment only, and the remainder receive a combination of work experience and academic enrichment. Their participants are predominantly in-school youth and, therefore, are largely between the ages of 14 and 17. Only 15% of the youth served by this program are between the ages of 18 and 21.

The largest provider offers predominantly a combination of work experience and academic enrichment. This provider coordinates with the school districts in the area so that its participants can earn academic units toward their high school graduation. Other providers focus on private sector internships, and have strong connections with employers. Two smaller program focus on out-of-school youth, combining GED training, or leadership training, with afternoon work site experience.

Academic enrichment is defined by this SDA, as with other SDAs, very broadly. Included under the rubric of enrichment are basic skills remediation, ESL, SCANS pre-employment skill development, as well as leadership and citizenship training. Further, the SDA clearly separates academic enrichment from work experience, although they are attempting to integrate these two facets of the Summer Youth program better than is currently the case.

The goals of the program are varied. The SDA would like to help participants prevent summer learning loss, but they also hope to increase the participants' confidence,

exposure to work concepts, contact with caring adults, and exposure to work environments and expectations. More broadly, the goal is to keep young people off the streets and in school during the summer months.

This SDA uses a slightly modified version of the CASAS as an assessment of work maturity skills and basic skills. These measures are standardized across all providers contracting with this SDA and are given as both a pre- and post-test to assess the impact of the program. Additionally, participants are given the Washington Basic Skills Competency System Assessment for reading and math, as a pre-test only, to identify deficiencies. Finally, some providers for this SDA have youth complete a portfolio as part of their assessment.

This SDA believes that the CASAS system is a useful one for assessing their overall performance. Primarily this is because it is grounded in work concepts and, therefore, students are less resistant to concentrating on the test. Also, the participants seem to like the competency-based elements of the system, according to the SDA. Specifically, the participants are appreciative of the definitive task definitions which are an element of the CASAS system. This clarity in expectation helps participants to understand what, exactly, is expected of them in order to be considered to have achieved a competency.

With this SDA, as with some others, there are no obvious ways to assess some of the more nebulous goals of the program. For example, it is unclear how one would measure confidence, or contact with caring adults and, even if such a measure were used, whether this would indicate real changes in the youths' opportunities or resources. As such, there is no clear way the SDA can document its performance in this area. They can, however, document whether they kept young people off the street by recording attendance data. Similarly, they can document exposure to the work world, as well as participants' various work maturity or basic skills, through the use of the CASAS. As a result, this SDA would seem to have an assessment system in place which is closely related to its primary goals.

Additionally, however, the SDA expressed interest in developing standardized measures for use with the Summer Youth program, but cautioned that such measures would also need to be flexible to allow creativity on the part of SDAs and providers. The CASAS system used by this SDA has some of these properties, but it also has the disadvantage of being too large, or too broad, for the Summer Youth program. The CASAS system is intended to be used with a minimum class-time of 120 hours. Because the Seattle summer program is only six weeks, and only part of the time is spent in academic enrichment, participants often do not receive such extensive classroom training. As a result, the CASAS system has some significant drawbacks as an accountability measure. The potential to develop an alternative system which is both standardized and flexible is discussed in greater detail below.

Finally, this SDA, and several of its providers, believe that portfolios offer an in-depth means to assess participants' progress during the summer. Indeed, the providers for this SDA use a variety of "authentic assessments," or project-based products, as means of

assessment. The potential of these authentic assessments to be useful in a national accountability system is discussed below.

Analysis of Client-Level Data From the Pilot Test. We received client-level data for 318 participants from this summer's pilot test. Each of these participants received at least some academic enrichment; some only received enrichment, and other received both enrichment and work experience. These data were received from 17 separate providers, or SDAs who serve as their own provider. The average number of respondents per provider therefore was 18.7. This average reflects the fact that we sampled at most 25 respondents from any one provider, and that many providers served fewer than this number. It is important to note that the selection of SDAs to participate in the pilot test was not made randomly, nor can we be assured that providers chose their own participant samples randomly. As a result, then, this average, and the results of our analyses which follow, should not be taken to be indicative of all SDAs or providers nationwide. Indeed, this study does not intend to generalize the results of statistical analyses to Summer Youth programs not included in the study. Instead, the results for these statistical analyses should be taken to be descriptive of those SDAs in our sample only.

We begin this section with a description of the quality of the data received from providers and SDAs. Next, we provide an overview of the types of assessment instruments used by the providers who forwarded their data to us. We then move on to discuss the results of our analysis, including those for reading skills, math skills, and pre-employment/work maturity skills. Finally, we discuss whether any particular measurement instrument makes it more likely that a participant will demonstrate a gain in a specific skill area.

Data transmission ranged from computerized submissions of database entries to mailed copies of long narrative evaluations. For each of these individual records, we identified, to the extent possible, the type of assessment instrument used, whether or not there were complete pre- and post-test scores, and what those scores were. These data were then entered into a statistical software file to be analyzed, as described below.

As may be expected in a pilot test, data quality ranged from those submissions which were quite accurate and complete, to those which were fraught with problems and were exceedingly difficult to translate into meaningful evaluation scores. As a result, for any particular analysis, our results are based upon somewhat fewer respondents than the overall total of 318.

The predominant test used for the participants in our sample was the Test of Adult Basic Education (TABE). Nearly half the participants (44.7%) received this test (N=142). Next most common were various components of the Comprehensive Adult Student Assessment System (CASAS). Approximately one-third of our sample participants (32.4%) were assessed using this instrument (N=103). Finally, a sizable percentage of the sample participants (16.4%, N=52) were assessed using the MAST measure. The remaining participants in our sample were assessed using a variety of instruments specific to the particular provider (i.e., Adult Basic Learning Exam). For only 5 of the participants was it impossible to determine which measure had been used for

assessment.

Because these measures assess participants somewhat differently, and produce scores which can not necessarily be easily compared, our analysis does not assess the specific level at which participants are functioning. Instead, we calculate difference scores using the pre- and post-test ratings, and assess the extent to which participants demonstrate gains over the course of the summer. Specifically, we subtract the pre-test score from the post-test score. As a result, a positive score indicates the student increased their performance during the summer; a negative score indicates a decrease; and a zero score means no change in performance. Thus, the analysis is restricted to those participants for whom we have complete pre-test and post-test scores using the same assessment instrument.

Of the 318 participants in our original sample, only 151 had both pre- and post-test scores. In some cases, this was because we did not receive data from the pre-test. This may be due to the fact that a few providers were informed after their programs began that they would be participating in the pilot test. As a result, their pre-test data may not have been as easily accessible as those providers who were aware early on that they were participating in this pilot study. In many more cases, however, the participants received a pre-test, but for a variety of reasons (ranging from the participant being absent on the day of post-test administration to lack of time among provider staff) received no post-test measure. In a few cases, participants were given a pre-test using one assessment instrument, and a post-test using a separate instrument. Because there is no standardized way to compare the two scores, these participants also had to be dropped from the analysis.

We begin our analysis by assessing the reading, or verbal, skills of those in our sample.³ For the purposes of this analysis, a participant “increased” their skills if their post-test score was higher than their pre-test score. A participant maintained their skills in this area if the change in their scores from pre-test to post-test was exactly zero. A decline in participants’ reading skills was indicated by a lower post-test as opposed to pre-test score. The results for this analysis are shown in Table 1. As can be seen in this table, of the 151 youth who had both pre- and post-test scores using the same assessment instrument, nearly three in four (N=109) either maintained their reading skills or improved them over the course of the summer.

³ A sizable percentage of providers assessed youth in two major categories: reading and mathematics proficiency. As a result, these are the two areas of greatest focus in this section. There were, however, some important variations. For example, some providers assessed what they label verbal skills, as opposed to reading skills. Where this occurred, we used verbal skills as a proxy for reading skills. Additionally, some providers forwarded to us scores for pre-employment/work maturity. Results for these analyses are described in the text.

Whether all of these skills are included in one’s definition of academic enrichment is not of issue in this section. Instead, we analyze results from each of these areas, to the extent possible, and take them to be meaningful in obtaining an overall view of academic enrichment. The reason for this approach is quite simple. If SDAs, or their providers, view such skills as part of academic enrichment, then results for these skills must be considered when discussing academic enrichment.

Table 1

Change in Reading/Verbal Scores		
	<u>Number</u>	<u>Percentage</u>
Increase in Scores	51	33.8%
No Change	58	38.4
Decrease in Scores	42	27.8

Note: Participants were included as having increased their score if their post-test value exceeded their pre-test value. No change indicates the pre-test score was identical to the post-test score. A decrease in score indicates the pre-test score exceeded the post-test score. The average gain across all participants was .65 grade levels.

Thus, a strong majority either maintained or increased their reading skills. For a program which lasts from six to eight weeks during a time when youth traditionally exhibit decreases in basic skills (the “summer learning loss” phenomenon) to demonstrate such results in part reflects the positive contributions made by Summer Youth SDAs. It should be noted, however, that the average gain was less than one academic year. In other words, while slightly more than one-third of participants achieved some gains in their reading skills, these gains were quite modest in magnitude.

One point which should be made about these data is that there was quite minimal variation between pre- and post-tests. It is striking that over one-third of all participants received identical scores on their pre- and post-tests. It is unclear to what this should be attributed, but it should be noted that previous national samples have arrived at very different conclusions using similar analyses. For example, one recently completed national study reported that only 16% of youth demonstrated no gain in reading scores, and only 19% of youth decreased in their scores.⁴ One explanation, then, for the results observed for this pilot study is that they are due to a non-random sample combined with a relatively small sample size. Regardless, we do not attempt to treat these data as generalizable to all SDAs nationwide. Instead, we use them only to support arguments made below concerning expectations of participant progress.

Next we consider the results for mathematics. This analysis was conducted similar to that for reading skills. The results are shown in Table 2. For this analysis, there were 140 participants for whom we had complete pre- and post-test results. Of these 140, slightly fewer than 80% (N=109) either maintained or increased their skills during the

⁴ D’Amico, Dickinson, Kogan, et al., *Evaluation of the Educational Component of the Summer Youth Employment and Training Program*. Menlo Park, CA. 1998. The sample was not randomly drawn and thus cannot be said to be generalizable to all participants in academic enrichment taken as a whole. Instead, respondents to the survey were drawn from students in Summer Youth programs that were purposively drawn to reflect a broad mix of enrichment activities.

course of the summer.

Table 2

Change in Math Scores

	<u>Number</u>	<u>Percentage</u>
Increase in Scores	67	47.9%
No Change	42	30.0
Decrease in Scores	31	22.1

Note: Participants were included as having increased their score if their post-test value exceeded their pre-test value. No change indicates the pre-test score was identical to the post-test score. A decrease in score indicates the pre-test score exceeded the post-test score. The average gain across all participants was .51 grade levels.

Once again, however, the gains were not large. Indeed, the gains demonstrated by participants in math skills were smaller than those for reading skills. While gains of a half grade during the course of the summer may seem impressive on the surface, it is instructive to remember that such gains often are the result of answering one additional question (out of 40) correctly in a post-test compared to the pre-test. Clearly, then, it would be a mistake to view these impressive gains as indicative that all SDAs should be expected to drastically increase their participants' basic skill scores. Such expectations are discussed in greater detail below.

These results for math scores also do not match those found by other recent studies. While it is similarly unclear why there are differences between the results for this pilot study and for those of other studies, it seems safe to assume that the small, non-random sample in this study affected the results in unpredictable ways. As a result, again we do not attempt to generalize these results to all other SDAs. Instead, we only take these results as descriptive of our sample, and use them to inform our analysis of the sample SDAs in later arguments.

Turning now to pre-employment/work maturity skills, there are far fewer participants in our sample for whom we have data. In large part this is because fewer than half the providers in our sample forwarded to us data which assessed these skills. As a result, we have complete pre- and post-test data on pre-employment/work maturity skills for only 66 participants.

One potential reason for the smaller sample size for pre-employment/work maturity skills is that not all SDAs view these skills as part of academic enrichment. As a result, they did not include them in this pilot study of measures for academic enrichment. Further, the assessment of pre-employment/work maturity is less straightforward than is the assessment of basic skills described above. Some SDAs ask employers to complete checklists with simple yes-no options for whether the participant has the skills. Often, this is done only at the end of the summer, making comparisons of pre- and post-tests impossible. Other SDAs, however, ask employers to complete checklists at both the

beginning and toward the end of a participant's work experience. Generally, the employer ranks the employee on a scale from one to five, with one being inadequate, to five, which represents an excellent rating. It is predominantly this form of checklist which is included in our analysis of pre-employment/work maturity.

Of the 66 participants for whom we have complete data, nearly all (95.5%, N=63) demonstrated a gain in at least one competency area. Somewhat fewer participants realized gains in two areas (77.3%, N=51). And very few participants realized gains in five of the 11 pre-employment/work maturity skill areas (12.1%, N=8). This is important because, under Title II-C guidelines, a youth obtains a pre-employment/work maturity competency only when s/he realizes gains in at least five skill areas. As we discuss below, then, if YECs are adopted for the Summer Youth accountability system, at the very least it seems that the requirements for competencies must be made less stringent to adjust for the relatively brief duration of the summer program.

One important consideration in our analysis of these client-level data is the extent to which the type of test affects the results. The reason this is important is clear: if it is more difficult to demonstrate gains on one test than it is on another, then one must adjust for the type of test when evaluating the performance of a particular SDA or provider. Only by adjusting for these variations can one be certain that difference scores of the individual tests are comparable. Fortunately, however, there were no effects for type of test in this pilot study. To assess this, we calculated the difference scores described above (post-test minus pre-test). We then examined these difference scores for the various test types in our sample (i.e., TABE, CASAS) to determine if one test led to more positive or negative scores.⁵ There was no effect for any of the tests, nor for any of the skill areas being assessed, indicating that each measure allows participants the same opportunity to demonstrate gains or losses.

We must be careful in overstating this result, however. What this test demonstrates is that, given our limited sample size, we cannot say that the differences in means between tests is statistically significant. With a larger sample size, or more precise measurements, though, our conclusions may be different. In other words, our relatively simple analysis suggests that one type of test does not lead to greater increases or decreases in scores than does any other. This analysis, however, is quite limited in power and, as a result, must be taken to be tentative.

These results offer a relatively straightforward analysis of the gains in basic and work maturity skills made by the participants in our sample. They also demonstrate, within the constraints described above, that each measure used by providers or SDAs is comparable in terms of its ability to observe gains or losses in these skills. However, in no way do these results demonstrate whether the various measures in our sample are well-suited to assess the goals or objectives of the SDAs or providers. An analysis of the

⁵ For this analysis, we created mean difference scores for each of the various types of tests. After a transformation to ensure that the variations in scores across tests were similar, we then used t-tests to examine whether the difference in mean scores for any of the individual tests was significant. There was no significant difference between either type of assessment instrument, or type of skill area.

appropriateness of these measures to the goals of the Summer Youth program, or to the SDAs and providers in this pilot study, is detailed below.

Difficulties in Assessing Academic Enrichment. Several SDAs, including some in the pilot study as well as several in the focus group, indicated that their definition of what constitutes academic enrichment may extend beyond what it is that DOL intends for this term. Because the definition of academic enrichment is vague, several SDAs have used it as an umbrella concept which encapsulates a wide variety of training and instruction. As a result, academic enrichment now includes, depending upon the SDA, basic skills training (i.e., math, reading, writing), pre-employment/work maturity skills, life skills training, basic economics, teambuilding, problems on the job, safety, sexual harassment, and labor market information training, among many other things. Such broad definitions complicate the design of an accountability system, because there are a wide variety of concepts which must be accommodated by a single system. But these broad definitions also make it extremely difficult to know what is meant when an SDA refers to academic enrichment.

One solution, suggested a few SDAs, is for DOL to define more precisely what is intended by the term academic enrichment. While all SDAs acknowledged the utility of a vague definition, which allows them the freedom to adapt their training to participants' and their communities' needs, they asserted that a more concrete definition of enrichment would allow them to know what it is, exactly, that DOL would like them to include in their training. In this way, they could design their curricula to ensure they include those materials of utmost importance to DOL, and it would be measures of these items on which an accountability system would be based. Additionally, the SDAs could include other training depending on the specific needs of their participants or communities. Of course, each SDA has a unique population and environment in which they function; as a result, measures for these SDA-specific items would not be included in an accountability system.

This solution, then, suggests that there would be at least two tracks on which SDAs trained their participants. First, with a more precise definition of what is included in academic enrichment, SDAs could focus their assessment efforts on those areas, and be held accountable on these items. Second, there would be additional areas of training in which an SDA could focus, depending on the needs of their participants, or the business community in which they operate. The results of all training in these areas would not be part of an accountability system because this training would be unique to a particular SDA.

But most SDAs are quick to point out the disadvantages to this approach, however. One problem is that by narrowing the definition of academic enrichment, DOL would implicitly create a new, unnamed category which includes all other forms of training or emphasis. In an effort to make the standards of accountability clear, DOL would have narrowed its focus to only a small portion of what is now considered academic enrichment. This defeats the entire reason for keeping the definition broad and somewhat vague in the first place. A related concern with this solution is that by focusing exclusively on the narrower definition of academic enrichment, DOL would be ignoring

much of what it is that SDAs do during the Summer Youth program. To understand this, one must only consider an analogous case: imagine if a high school were to be held accountable only for their students' performance on a chemistry exam. Clearly, this exam is not reflective of the total effort made by the high school; instead it is a meaningful reflection of only one aspect of the school's efforts. Similarly, holding SDAs accountable for only part of what they do seems problematic.

A better alternative to the solution of more precisely defining academic enrichment may be to ensure that any accountability system put in place remains flexible. As mentioned, DOL has given the SDAs flexibility in defining academic enrichment. A highly rigid system which does not allow for any adaptation of assessment on their part therefore would not provide an accurate national picture of the variety of activities offered by Summer Youth SDAs. Creating additional standardization may be important, and relatively easily achieved by narrowing the definition of academic enrichment. But the majority of SDAs do not wish to conform their definitions of academic enrichment, or even competency gains, to a single, narrow one, and this is seen as a major hurdle for any accountability system. To the extent that an accountability system does not need to constrain grantees in this way, SDAs and providers seem willing to use measures to help document their performance and achievement.

The assessment system around which the pilot study was based was the YEC system. As discussed above, one advantage for using the YEC system is that it offers at least three different areas in which an SDA can focus, thereby allowing for some flexibility. This type of system could be useful for accountability if SDAs were allowed to choose from among the skill areas those most relevant to the SDA's training goals. If, for example, SDAs were allowed to choose from among the three skill areas in the YEC system, and be assessed on whichever of the areas is their focus, they would then be held accountable on the very indicators which are of greatest importance to their goals. But it is entirely plausible that many SDA's goals are not captured by the YEC system. As can be seen in the descriptions of each above, many SDA's goals are not compatible with the three skill areas contained in the YEC system. This is especially true when one considers that this pilot study, and presumably the entire accountability system, focuses only on those Summer Youth participants receiving academic enrichment, and not those receiving work experience. As such, it is even less likely some of the areas of the YEC system, particularly those involving job-specific skills, would be relevant to the Summer Youth SDAs.

It is important therefore to offer a wider array of potential measures than even the YEC system offers, from which SDAs can choose those most pertinent to their objectives. Potentially, then, DOL could adopt a menu approach to an accountability system for the Summer Youth program. Such an approach would identify several options for accountability measures – including indicators of pre-employment/work maturity, basic skills, job-specific skills, life skills, teambuilding, and customer satisfaction, among many potential others – and allow the SDA to choose from this menu those measures which best capture the goals of their program. Additionally, SDAs could choose those skill areas which are of greatest importance to the businesses in their communities. In

this way, the national accountability system would be able to assess the accomplishments of SDAs, nationally, by using measures which are highly relevant to each of the contributing SDAs. This approach is described in greater detail below, in the conclusion to this report.

Difficulties with the Duration of the Program. In addition to the confusion surrounding the definition of academic enrichment, and what aspects of this concept would be included in an accountability system, many SDAs and providers also are concerned that DOL will have unrealistic expectations in light of the brief duration of the Summer Youth program. Specifically, SDAs are concerned that the six to eight weeks in which participants are generally in the program do not allow sufficient time to implement sophisticated accountability measures. For example, some tests used by SDAs in this pilot test required two or three days on the part of participants to complete. Using such tests for both a pre- and a post-test removes an entire week from the activities offered to the participants. This represents more than 10% of the time for which the programs are offering training. Clearly, then, such measures would, in the interest of documenting performance, hamper the ability of participants to achieve gains in the very areas being documented.

Indeed, many SDAs indicated that the measures they use for their own assessment purposes were chosen at least in part because they require less time to complete than do other options. This may seem logical, but it is not always well understood when implementing accountability measures on a wide scale. Thus, it is important as a national accountability system is put into place to be cognizant of the need for brief measures to ensure that program quality is not reduced in an effort to assess SDA performance. One factor which partially alleviates this concern is that a national accountability system likely will not survey all SDAs, nor all participants within a particular SDA. Instead, both SDAs and youth likely will be sampled. As a result, the burden imposed on SDAs and their participants will be greatly minimized to the extent that DOL efficiently samples youth to be included in the accountability system.

Another concern with the short duration of the Summer Youth program is that no matter what assessment measure is used, the definition of a success, or competency, may have to be relatively modest in light of this short time frame. For example, if DOL chooses to implement the YEC system as an accountability system for Title II-B, as it has with the year-round Title II-C program, the Department likely would have to revise somewhat the definition of a success or competency. Specifically, the requirement under the YEC system in Title II-C that at least 2 of 3 of PE/WM skills, basic skills, or job-specific skills must be achieved before credit can be claimed seems unreasonable.

Moreover, it might be unreasonable to expect anyone to attain job-specific skills, or appreciably increase basic skills, given these same circumstances. Indeed, while many participants in our pilot study maintained their basic skills over the course of the summer, fewer than half demonstrated gains in these skills. And the gains which were observed were rarely large in magnitude. Thus, it seems unlikely that Summer Youth participants could reasonably be expected to demonstrate such gains in basic skills during such a short program. It must be mentioned again, however, that given the problem of summer

learning loss, helping participants to maintain their basic skills over the course of the summer could be taken as a noteworthy achievement.⁶

Finally, the requirement that competency must be achieved in at least 5 of the 11 PE/WM skill areas similarly seems too onerous. As was evidenced by the quantitative data discussed above, very few participants achieve competency in five of these skill areas during the brief summer program. Maintaining the rigorous standards of the Title II-C program therefore would likely ensure that the vast majority of Summer Youth SDAs would not meet their standards simply because the participants do not have adequate time to accomplish so many gains. Thus, for purposes of the Title II-B program, if YECs were implemented as an accountability measure, we suggest requiring somewhat less attainment than is required in the Title II-C program in order for a participant to be considered a success.

Using similar, but less stringent, YEC benchmarks would reflect the goals of many of the SDAs while acknowledging the difficulty in achieving large gains during such a short period. Additionally, as described above these benchmarks would acknowledge the diversity in academic enrichment offered by the various SDAs. For example, for those SDAs who concentrate their enrichment efforts on classroom instruction covering basic skills, their participants would be expected to demonstrate some competency attainment in these skills. But those SDAs who focus their enrichment efforts primarily on integrating academics into work experience, and attempting to provide youth with a positive employment experience, could document their performance by demonstrating their participants had attained some of the 11 pre-employment/work maturity skills. In this way, SDAs and their providers could continue to offer a diverse mix of programs to participants, and to emphasize somewhat different goals, while the national accountability system still would be able to capture a picture of the various achievements of Summer Youth programs.

The Fit Between SDAs' Goals and Assessment Measures. Two additional findings emerge when comparing across all SDAs. First, many of the SDAs in the pilot study relied heavily on the assessment of basic skills, using a variety of measures to assess their participants both before and after the program. However, very few of these SDAs indicate that increasing basic skills is an important objective of their program. Indeed, many SDAs actively avoid large doses of classroom instruction tailored to increasing basic skills, for fear of alienating their participants who typically have not performed well in academic settings. Further, many SDAs indicate that their participants have trouble making relatively simple life decisions and, therefore, they focus on increasing these skills instead of basic skills. Thus, there is a fundamental incongruency between the

⁶ For evidence concerning summer learning loss, see Cooper, Nye, Charlton, Lindsay, and Greathouse, "The effects of summer vacation on achievement test scores: a narrative and meta-analytic review," in *Review of Education Research* 66 (Fall, 1996). Among the evidence discussed, the authors note that disadvantaged youths' learning parallels that of other youth during the academic year. Over the course of the summer, however, disadvantaged youths' skills decline precipitously, unlike other youths' scores. This gap is then compounded each year during summer vacation. To the extent that this decline is prevented, then, disadvantaged youths' skills should parallel those of other youths.

measures used for assessment and the goals or objectives of the programs. Perhaps it is because of this fundamental incongruency that the gains observed in the quantitative analysis described above were so modest. SDAs hardly can be expected to impart large gains in basic skills to their participants when this is not their focus, and the results from our pilot study confirm this assertion.

Results from a recent customer satisfaction survey of Summer Youth participants also support this conclusion.⁷ For example, of those Summer Youth participants who received academic enrichment services, only 44% agreed that the program was “quite” or “very” helpful in improving their reading/writing. Similarly, only 45% agreed it was quite or very helpful in improving their math skills. This reflects the fact, once again, that most SDAs do not emphasize gains in these basic skills in their summer programs.

To the extent that these basic skills are what DOL intends by academic enrichment, then an accountability system for such enrichment seems destined to have significant obstacles. Because many SDAs do not focus heavily on these skills in their Summer Youth programs and, as a result, not many gains are demonstrated, establishing an accountability system based on these skills cannot be expected to indicate great success on the part of SDAs. This clearly would be an ineffective system, unless SDAs changed their approach to focus almost exclusively on training participants in their basic skills. While this may be an option, it seems like an example of putting the cart before the horse. Rather than establishing an accountability system which documents what it is that SDAs do, this alternative would almost force SDAs to alter their programs in an effort to meet their accountability goals.

If, on the other hand, the definition of academic enrichment encapsulates a great deal more than these basic skills, then SDAs have greater flexibility in their programmatic objectives. Given the wide variety of objectives across the SDAs, those in the pilot as well as those in the focus group, this seems to be a more suitable alternative. Here, SDAs would continue what it is they are doing, and the assessment system would be designed around these various goals. In this way, the accountability system avoids the perverse incentive of asking SDAs to fundamentally alter their programs in an effort to meet accountability needs.

A second finding which emerges in comparing across programs is that instead of the basic skills which they are assessing, many SDAs emphasize “softer” objectives (e.g., a feeling of “I can,” increasing self-confidence, etc.). This, too, can be seen in the results from the recent customer satisfaction survey. For example, 66% of participants reported the program to be quite or very helpful in improving their self-confidence. Further, 74% agreed that the program was quite or very helpful in helping them to set goals for the future. Also, fully 80% agreed that the Summer Youth program was quite or very helpful in teaching them about work or careers. Clearly, then, according to their own ratings,

⁷ D’Amico, Dickinson, Kogan, et al., *Evaluation of the Educational Component of the Summer Youth Employment and Training Program*. Menlo Park, CA. 1998. As described above, the sample was not randomly drawn and thus cannot be said to be generalizable to all participants in academic enrichment taken as a whole.

more participants in the program received help in areas other than basic skill instruction. These data reflect very directly the objectives of the SDAs.

Customer satisfaction measures related to SDA objectives may, then, be a legitimate alternative in attempting to capture the goals of the SDAs. Having asserted this, however, many of the SDAs' goals are exceedingly difficult to quantify or measure objectively. While participants certainly can indicate their satisfaction with a program's instruction or work opportunities, it is less clear how to define and measure whether they gained a sense of "I can," or whether they gained "useful life skills." The creation of a useful accountability system is therefore complicated by, on the one hand, the fact that readily measurable, standardized indices (of basic skills, for example) do not capture the goals of the programs, and, on the other hand, the fact that those goals are extremely difficult to measure in any meaningful way. Whatever accountability system is put in place, then, must strike a balance between these two complications.

Whom To Include in an Accountability System. An additional consideration in an accountability system is whether to include all Summer Youth participants in this system, or only some subset of these participants. As mentioned above, consistent with DOL's request, only those participants receiving academic enrichment (either alone, or in combination with work experience) were included in this pilot test. However, for the national accountability system, we believe there are two possible approaches which DOL can consider. First, DOL can define academic enrichment much more narrowly, as described above, and measure it accordingly. Alternatively, DOL can keep the definition of enrichment broad, but if this is the case, we suggest expanding the accountability system to take into account the experiences of youth receiving SYETP traditional work experience. Based on the findings of this pilot study, as well as additional research, there are several reasons why we advocate this position.

First, according to a recent national study of the Summer Youth program,⁸ more participants in Title II-B receive just work experience than receive academic enrichment or education (either alone or in combination with work experience). While this is not true for the SDAs in our sample, it should be kept in mind that our sample was not randomly selected, nor were the providers from whom we obtained much information. As a result, results from the recent national study should be taken as more indicative of the general experiences of Summer Youth SDAs. Given this, it would be unfortunate if the national accountability system for the program ignored the experiences and accomplishments of over one-half the program's participants.

Additionally, DOL has vigorously argued that SDAs should ensure that work experience assignments are high quality learning experiences that "build and refine a strong foundation of work, employment competencies, and the discipline of work" (language taken from the 1996 TEG). Indeed, many of the SDAs in this pilot study went to great lengths to describe the quality of their worksite placements. Including work experience in the accountability system would further serve to demonstrate DOL's

⁸ Westat, Inc., *A Study of the JTPA Title IIB Program During the Summer of 1993: Final Report*. Westat, Rockville, MD. 1994.

commitment to high quality job placements, and would provide a way of demonstrating the value of such placements.

Also, as has been made clear above, much of what is labeled by SDAs as academic enrichment focuses on pre-employment/work maturity, or other “softer” skills. These skills are precisely the major objectives associated with traditional work experience. As a result, participants in academic enrichment are being trained in skills similar to those offered to participants in work experience. An accountability system which assesses these skills only for those participants receiving enrichment captures only one segment of youths’ receiving training in them. Such an accountability system may well underestimate the overall impact of the program.

Moreover, as demonstrated by the descriptions of each of the SDAs in the pilot test, and by the results described above, definitions of what academic enrichment should or does entail vary considerably. Thus, an activity that one SDA labels “enrichment,” may not be so labeled by another SDA. As a result, participants receiving the same training may be included in accountability measures by one SDA and not by another. Including all participants in the accountability system makes these distinctions irrelevant.

Finally, many young people (particularly those receiving both enrichment and work experience) receive enrichment for only 10 or 20 hours during the entire course of the summer, while, in contrast, they receive 100 or more hours in work experience. It is unclear that one should include participants receiving such a trivial dose of “academic” intervention in an accountability system meant to measure only the gains from enrichment. Excluding them means one is making an arbitrary cutoff (e.g., excluding those who receive fewer than 50 hours of enrichment?). Capturing all participants in the accountability system obviates this difficulty.

These considerations are not meant to understate the challenges of including work experience participants in the accountability system. Most notably, it would be much more difficult to conduct assessments in myriad work sites in which just a few participants might be placed than to conduct assessment in a classroom setting, where at least most academic enrichment participants will be served. Also, it is not clear that work experience assignments currently are always associated with clear learning objectives, on which the assessment of learning gains could be based.

But the advantages of including them are substantial enough that it is at least worth considering the issue as DOL moves forward toward a final accountability system. Moreover, since the national accountability system will doubtless rely on sampling, not all work experience slots but just a sample of them will be affected, thereby limiting the logistical challenges. Given that the results from this pilot study indicate that many SDAs have “softer” goals for their participants regardless of whether these participants receive academic enrichment or work experience, an accountability system trying to assess performance on these goals likely should include both sets of participants in its sample.

Limitations of the YEC System: Alternatives. This summer’s pilot test of an accountability measure for the Title II-B assessment system used various adaptations of YECs as a starting point for very good reasons. As mentioned, these adaptations of YECs

give SDAs substantial flexibility, providing a wide umbrella in which very diverse programmatic activities could be covered. Additionally, SDAs already had a YEC assessment system in place for their Title II-C programs, and were well familiar with DOL's requirements for a "sufficiently developed system."

But YECs also have certain important limitations in providing an accurate portrayal of the program's chief accomplishments. Several of these already have been pointed out by SDAs above. For example, the pilot test intended for the YEC system to capture achievements vis a vis academic enrichment, but arguably some of the competencies covered by this assessment system do not mesh very well with DOL's definition of what academic enrichment should entail. For example, the latest TEGL defines academic enrichment as participation:

"in a structured learning experience (on or off the job) where SCANS foundation skills and competencies (or reasonable variations thereof) and/or other academic disciplines are taught and progress can be measured and documented."

Several of the PE/WM skills areas defined in the YEC system (e.g., such as using labor market information, preparing resumes, filling out applications, developing interviewing skills, among others) only by the wildest stretch could be considered teaching "SCANS skills and competencies or reasonable variations thereof." In short, YECs assess skills in areas that do not seem to be implied by the definition of what academic enrichment should entail.

However, many of these skills are included by SDAs in their definition of academic enrichment. As a result, while the YECs do not seem well-suited to measuring what DOL intends as academic enrichment, they may be able to measure reasonably well those skills which SDAs define as academic enrichment. As described above, therefore, if DOL wishes to implement YECs for the national accountability system, it is incumbent upon them to modify their definition of academic enrichment. Specifically, DOL would have to broaden the definition of enrichment so that it encompasses those skills assessed by the YECs. Using YECs without a modification of the definition of enrichment would send very conflicting and confusing messages to the field as to what constitutes or should constitute academic enrichment. In fact, the goals of SYETP traditional work experience seem to mesh better with the YEC assessment system than does academic enrichment as it currently is defined.

As discussed above, YECs assess some skills not covered by the definition of academic enrichment. But, conversely, the definition of academic enrichment covers skills development in areas not encompassed by the YEC assessment system. For example, efforts to impart thinking skills or problem-solving skills, specifically a part of the SCANS framework, are not covered by most YEC assessment systems, nor are efforts to "teach other academic subjects." Thus, by using the YEC assessment system, some of the programs' chief accomplishments may be missed entirely. Indeed, several SDAs in our pilot test indicated that the measures they used did not adequately capture the overall goals of their program. This is because academic enrichment, as defined by DOL, and many of the SDAs, includes concepts not covered by the YECs and, therefore, using

YECs as an accountability measure will capture only some aspects of the participants' competencies achieved through the Summer Youth program.

Finally, an evaluation of the YEC system completed in 1992 identified a number of problems with YEC implementation.

⁹ Most importantly, it was discovered that the quality of YEC programs varied considerably and that some SDAs set low levels for the attainment of competencies. Moreover, their assessment systems varied wildly, and many lacked rigor. By implication, the definition of what constitutes the attainment of a competency can be expected to differ markedly from one SDA to the next, and, as noted in this report, often these definitions differ greatly even from one provider to the next within the same SDA. Given these circumstances, evidence that a youth attained a competency could mean very little in practice.

Thus, the YEC system used in this pilot study does not offer great advantages for a national accountability system. Although many SDAs are familiar with it because of its use in the Title II-C program, the many drawbacks described above seem to outweigh this general benefit. The greatest difficulty in using YECs for a national accountability system at this point is that they assess participants in many areas which do not fit within DOL's current definition of academic enrichment and, conversely, they do not assess some areas which do seem to be included in the definition. Thus, for YECs to be a highly useful assessment instrument, either they or the definition of academic enrichment must be modified so that there is a stronger relationship between the two.

Upon concluding there was a poor fit between YECs and academic enrichment, we began to explore other potential alternatives for accountability measures. One option offered by several SDAs is to use authentic assessments for the participants of the Summer Youth program. Authentic assessments are those assessment instruments which are a compilation of youths' actual work, rather than their performance on standardized tests. For example, many SDAs indicated that the portfolios produced by their youth were the best indicator of the overall quality of the participants' training and experience in the Summer Youth program. These portfolios generally included a resume, a cover letter, letters of reference, and examples of the youths' work in terms of writing, reading, mathematics, problem-solving, and cooperation. By the end of the summer, then, participants had developed a useful portfolio which contained samples of their work throughout the program. These samples offer an indication of the skills the youth has, and provide potential employers with a wealth of information on which to assess the participant.

Despite the apparent advantages of the authentic assessment approach, there are several reasons why it may not be suitable for a national accountability system. First, each portfolio (or similar authentic assessment measure) contains several examples of one's work. While this offers great depth in examining a single participant's progress, it

⁹ Dickinson, Means, Kelley, 1992, *Youth Employment Competency Programs in JTPA*. SRI International.

is quite cumbersome to evaluate large samples of participants because of the sheer volume of information included in each portfolio. And attempting to coordinate data collection using these types of assessments on a national level seems virtually impossible or, at the very least, too cumbersome to be useful.

A second drawback to portfolio assessment approach is that each participant's work will be somewhat different from any other participant's. Where this is the case, it is unclear how to evaluate the achievements or competencies of these participants consistently. As a result, using portfolio assessment on a national level would involve highly subjective evaluations, and one would be forced to make arbitrary decisions concerning what represents a gain or competency in a particular portfolio. Clearly, such subjectivity and arbitrariness would undermine the utility of an accountability system. While the advantages offered by the use of portfolios - including greater depth and a broader range of observable skills - may outweigh their subjectivity for an individual employer, therefore, on a national level this subjectivity makes the use of portfolios a poor choice for accountability. Despite the ascribed utility of such assessments, then, the use of portfolios seems an unappealing alternative in the effort to establish a national accountability system.

Another potential alternative suggested by many SDAs focused less on testing of specific skills. Several SDAs indicated that one measure of success for their program are simple counts of participant attendance, and the rate of return to full-time school, either high school or post-secondary. The argument for this is that one can not expect gains in academic or work skills during such a short program, but that the programs do affect youths' decisions about school and career by providing a positive experience for them. To the extent that this is true, youth should be more likely to return to full-time school, and should be likely to attend each day during the summer program. DOL could, therefore, establish a national system of accountability which documents attendance and rate-of-return to full-time school.

To the extent that DOL intends for SDAs to conduct some specific training, this alternative is ill-suited for accountability measures. This is because the measures do not assess any training whatsoever. Instead, they measure whether youth were interested enough to show up for the summer program, and interested enough to return to school. Without some specific measures assessing the training DOL considers important for the Summer Youth program, SDAs will not be accountable for the training. Thus, attendance data and return-to-school may be useful aspects of an accountability system, but only in concert with assessments of the skills gained or maintained as a result of the training provided by SDAs.

Conclusion.

This report has summarized the results of a pilot test conducted in the summer of 1998 of accountability measures for the Title II-B Summer Youth program. Eight SDAs provided information concerning how they had adapted the YEC system to document the performance of their programs. In addition to extensive discussions with these SDAs, their providers were contacted, and many provided access to their client-level data. Also,

a focus group was conducted in an effort to solicit the opinions of several SDAs not participating in the pilot study.

Overall, the SDAs had little trouble implementing the assessment measures described in the study. In large part, this is because the study sought to document the measures each currently had in place, and to determine how well those measures captured the goals of the individual SDAs. It is clear from this study that SDAs for the Summer Youth program can successfully implement accountability measures. Of course, if DOL chooses to adopt measures unfamiliar to SDAs, or requiring significantly more effort than those already in place, it is not certain that SDAs will be readily able to adapt. It is recommended, therefore, that as DOL moves toward putting in place a final accountability system, that it either utilize measures already familiar to SDAs or, alternatively, that it work closely with several SDAs so that they know what is expected of them as they move to adopt the new measures. In this way, too, DOL will better be able to meet the needs of the SDAs offering services as part of the Summer Youth program.

This pilot study made use of the YEC system for its accountability measures. As mentioned, this was done to ensure that SDAs had extensive familiarity with the measures. It was learned through this pilot study that the use of YECs has several drawbacks which outweigh their advantage of being familiar to SDAs. As a result, it is recommended that either DOL choose an alternative system of accountability for the summer program, or that the YEC system be significantly modified in an effort to fit better with the purposes of the accountability system. Specifically, because DOL wants to establish an accountability system for participants in academic enrichment, it is recommended that any modification of the YEC system be done to ensure that key components of academic enrichment are covered, and items extraneous to academic enrichment not be included in the accountability system. The current structure of YECs fails to accomplish either of these goals successfully.

We also recommend that DOL reconsider the definition of academic enrichment. Purposefully left vague to encourage SDAs to remain flexible in their program activities, the definition has become instead a source of great confusion. SDAs include a wide variety of activities into their definitions of academic enrichment and then are left uncertain about which of these is important to assess in their participants. As a compromise, many measure their participants' basic skills, with little or no intention of seeing meaningful change in the scores from the beginning to the end of the summer. Many also indicate that the areas in which they most strongly influence participants are not being assessed and, as a result, the true successes of the summer program are undocumented. In large part, this incompatibility between the goals of summer programs and the assessments they use stems from the confusion over the proper definition of academic enrichment.

Two alternatives were identified for refining the definition of enrichment: one option is to narrow the concept; the other is to leave it broad, but make it more precise. We recommend the latter option be chosen because it ensures that SDAs have sufficient flexibility to adapt to the needs of their participants, or the communities in which they

operate, but it removes the rampant confusion surrounding the term. As such, we recommend that DOL identify which specific skills, or areas of skills, are included in the concept of enrichment, and include them explicitly in the definition. For example, the current definition of enrichment includes training in “SCANS competencies and/or other academic disciplines.” While both of these areas are important, many SDAs also incorporate life skills training, health and safety training, and basic economics, among other areas, into their definitions of enrichment. We recommend that DOL determine whether these types of training are adequately labeled as enrichment and, if so, include them as options in the definition. In this way, the definition still allows for sufficient flexibility, but offers precise explanations of what areas can be considered enrichment.

Combined with this revision of the definition of academic enrichment, we recommend that DOL consider adopting a menu approach to accountability measures. By this we mean that DOL could provide a list of acceptable accountability measures for enrichment, preferably sorted by the wide variety of skill area training (i.e. pre-employment/work maturity, basic skills, life skills, etc.) currently offered by SDAs. SDAs would then be allowed to choose from among this list those areas in which they concentrate their efforts. This would allow the SDA to continue providing a wide array of activities without having to assess their participants on activities which are not emphasized in the program. This makes the results more informative to the SDA. Also, because participants can more easily observe the relationship between what they are learning and what is being assessed, the process of assessment becomes more intrinsically interesting to them. And, in a national accountability system, DOL would be able to assess quickly which areas were receiving the most emphasis because SDAs would choose from those areas for their accountability measures.

Potentially, in this menu-based system, there could be some accountability measures which all SDAs would be asked to adopt. These may include relatively simple measures of basic skills or pre-employment/work maturity skills. This would be useful if it is these areas on which DOL wants SDAs to focus. If this is the case, then by mandating these areas as necessary for SDAs to assess, DOL would communicate that these areas should be emphasized in the Summer Youth program. For their part, there would be greater clarity in expectation for SDAs, and they still could assess in those areas of unique emphasis by choosing those areas as assessment measures for their program. Of course, if DOL does not deem it necessary to require any measures of SDAs, then they would have entire flexibility to choose their accountability measures to best fit their programmatic objectives.

Regardless of whether there are mandatory assessment measures in a menu-based system, DOL should remain cognizant that the Summer Youth program is short in duration and, as a result, large gains in any type of skills are unlikely. Thus, as a national accountability system is put in place, the goals of that system must remain relatively modest. For example, if basic skills among youth are assessed, a successful participant may simply maintain their initial skills rather than demonstrate gains. In other words, the avoidance of summer learning loss could be considered a success. Similarly, if life skills are being assessed, the acquisition of one or two skills may be enough to be considered a

success. This is not meant to lower standards of performance or training. Instead, it is meant to acknowledge that somewhat fewer gains are likely to be seen in a six to eight week program, especially one which has as diverse a set of objectives as the Summer Youth program, than are likely to be seen in longer and more extensive programs (for example, Title II-C).

This study has also documented the relative utility of customer satisfaction as an accountability measure. While DOL may desire to assess program performance on a broader array of measures than these, satisfaction does indicate that participants received something positive from the program. Further, many SDAs have as a primary goal engendering within their participants a feeling of self-confidence and satisfaction with their efforts. Customer satisfaction measures offer an excellent way in which to assess performance on these goals. Therefore, if DOL adopts a menu-based approach to accountability measures, customer satisfaction could be one area which SDAs could choose to be assessed.

Finally, we recommend that DOL consider implementing an accountability system which assesses all Summer Youth participants, not simply those who receive academic enrichment. The reasons for this are specified in detail above, but they include the fact that a majority of participants receive more hours of work experience than they receive of academic enrichment. Further, many SDAs integrate their enrichment activities with their work experience offerings, and the goals for participants often are similar in each type of experience. With this in mind, assessments of academic enrichment would only tell, at most, half the story. Because a national accountability system will certainly rely on sampling, the drawbacks to including work experience participants are significantly ameliorated.